|  |  |
| --- | --- |
| **Order** | Generally contingent, i.e. results of earlier steps are necessary for subsequent steps. This is also the reason why trip linking and mode revision were added to Rulesy—they’re improved with fixes applied first, and subsequent steps also require them. |
| **Db Functions** | Spatial functions, regular expressions, and datetime functions are all available as part of any query clause—they can be combined however we want. This allows a lot of power for specifying conditions or manipulating the data. |
| **General Approach** | Apply solutions where supported by the bulk of evidence, but leave a revision code and keep the original tripid to trace what was done (i.e., manual revision may alter Rulesy edits). |
| **Continual improvement** | Glad to enhance or add to what’s there; again, the above tools create plenty of options.  I’ll spend more time investigating patterns to correct; whatever insight Toolsie can add, will also get incorporated. |
| **Intent: Feed Toolsie/Fixie** | A view can provide all the fields in the format Toolsie prefers.  Rulesy not currently coded to publish the data—it is not the last stop. Some fields in question. |
| **Destiny: Elmer** | Easy import due to shared platform; haven’t yet coordinated with Chris on all aspects.  Additional geometric assignments can be easily added in Elmer; keeping mainly to reported fields and those necessary for cleaning. |

Rulesy overview:

STEP 0. **Settings independent of data tables**

1. Create and populate reference tables for mode categories used later
2. Create and populate table of NULL flags (new to 2019 data delivery)

STEP 1. **Load data from fixed format .csv files.**

1. Add additional fields, including geometry; create indices
2. Identify legitimate home location
3. Populate missing primary work or school locations if primary work or school destinations exist

STEP 2. **Set up auto-logging and recalculation**

1. Set up and enable auto-logging for field changes
2. Procedure to recalculate derived fields. (This is used both in Rulesy and Fixie.)

STEP 3. **Corrections to purpose, etc fields**

1. Update traveler counts when the number of household co-travelers is larger.
2. Origin purpose assignment for home and work
3. Destination purpose
   1. Home or work
   2. Revise return portion of single-stop loop trip
   3. Revise non-student ‘school’ purpose to pickup/dropoff, family activity, or personal business
   4. Fill missing purpose from hints in dest\_name
   5. Fill missing purpose from adult co-traveler who reported it
   6. Fill missing purpose from household member or other traveler to the same location, if consistent

STEP 4. **Trip linking**

1. Concatenate all modes into a single text field (similar procedure used for transit\_lines and transit\_systems).
2. Remove trip components into another table:

* prior record does not end at home or work
* trip purpose is either the same, or 'change modes'
* duration between legs is either less than 15 minutes, or if purpose 'change modes', less than 30 minutes.
* either there is a mode change, or both modes are transit

1. denote trip component which have aggregate properties that disqualify them as a single trip:

* two instances of pool start, parking, park & ride, change vehicles
* over 5 components
* non-adjacent repeating transit lines (i.e. loop)
* returning to the same geography as the prior trip

1. for legitimate linked trips, replace the initial component with the aggregate linked trip data
2. Recharacterize mode (including access/egress modes)

* Eliminate adjacent repeated values
* Characterize access and egress trips, separately for 1) transit trips and 2) auto trips. (Bike/Ped trips have no access/egress)
* Remove access/egress modes from 1) transit and 2) auto trip strings--not only at the ends, but also the middle.
* Split the concatenated field into separate mode fields

1. Split transit\_lines and transit\_systems into separate fields

STEP 5. **Mode number standardization, including access and egress characterization**

1. Eliminate repeated values for modes, transit\_systems, and transit\_lines
2. Characterize access and egress trips, separately for 1) transit trips and 2) auto trips. (Bike/Ped trips have no access/egress)
3. Remove access/egress modes from 1) transit and 2) auto trip strings--not only at the ends, but also the middle.
4. Split the concatenated field into separate mode fields (same for transit systems & transit lines)

STEP 6. **Harmonize passenger trips and purposes**

1. Insert trips for those who were reported as a passenger by another traveler but did not report the trip themselves
2. Recode driver flag when mistakenly applied to passengers and a hh driver is present
3. Recode work purpose when mistakenly applied to passengers and a hh worker is present
4. Add trips where destinations are more than 500m distant

STEP 7. **Revise travel times for excessive speed trips**

1. Adjust (preferentially departure, but also arrival if necessary) using results from Bing Travel Matrix API
2. Mode changed from non-vehicular to vehicular if travel window is too narrow and vehicular time approximates reported time

STEP 8. **Flag inconsistencies for further scrutiny**

1. Ends day, not home: destination is not home and purpose not 52, 55, or 97
2. Starts, not from home: initial diary trip is not home
3. (Destination) Purpose missing
4. Initial origin purpose missing
5. Mode\_1 missing
6. Lone trip: only one trip per person
7. Underage driver: driver = 1, age < 4
8. Unlicensed driver: driver = 1, license = 3
9. Non-worker + work trip: employment status > 4, dest\_purpose = 10, 11 or 14
10. Excessive speed: walking > 20mph, biking > 40mph, driving >85mph, other > 600mph
11. No activity time after: less than 1 minute between trips
12. Identical location as next trip: identical lat & lng
13. Time overlap: same person traveling two trips at same time
14. Same transit line listed 2x+
15. Purpose at odds w/ dest: home but not home purpose trip or vice versa; primary workplace but not work purpose trip or vice versa
16. Missing next/prior trip link: destination of prior trip not the same as origin of subsequent trip
17. “Change mode” purpose – i.e. code 60, not waiting at airport or ferry terminal
18. Too long at dest: limits for three groups of destination purposes
19. Non-student + school trip: not PT/FT student and over 17yrs